

# CONTENT INDEPENDENT DATA COMPRESSION METHOD AND SYSTEM

## CROSS-REFERENCE TO RELATED APPLICATIONS

- This application is a Continuation-In-Part of U.S. Patent Application Serial No. 09/705,446, filed on November 3, 2000, which is a Continuation of U.S. Patent Application Serial No. 09/210,491, filed on December 11, 1998, which is now U.S. Patent No. 6,195, 024, issued on February 27, 2001.

## BACKGROUND

### 1. Technical Field:

The present invention relates generally to a data compression and decompression and, more particularly, to systems and methods for data compression using content independent and content dependent data compression and decompression.

### 2. Description of Related Art:

Information may be represented in a variety of manners. Discrete information such as text and numbers are easily represented in digital data. This type of data representation is known as symbolic digital data. Symbolic digital data is thus an absolute representation of data such as a letter, figure, character, mark, machine code, or drawing,

Continuous information such as speech, music, audio, images and video, frequently exists in the natural world as analog information. As is well known to those skilled in the art, recent advances in very large scale integration (VLSI) digital computer technology have enabled both discrete and analog information to be represented with digital data. Continuous information represented as digital data is often referred to as diffuse data. Diffuse digital data is thus a representation of data that is of low information density and is typically not easily recognizable to humans in its native form.

There are many advantages associated with digital data representation. For instance, digital data is more readily processed, stored, and transmitted due to its inherently high noise immunity. In addition, the inclusion of redundancy in digital data representation enables error detection and/or correction. Error detection and/or

correction capabilities are dependent upon the amount and type of data redundancy, available error detection and correction processing, and extent of data corruption.

One outcome of digital data representation is the continuing need for increased capacity in data processing, storage, and transmittal. This is especially true for diffuse data where increases in fidelity and resolution create exponentially greater quantities of data. Data compression is widely used to reduce the amount of data required to process, transmit, or store a given quantity of information. In general, there are two types of data compression techniques that may be utilized either separately or jointly to encode/decode data: lossless and lossy data compression.

Lossy data compression techniques provide for an inexact representation of the original uncompressed data such that the decoded (or reconstructed) data differs from the original unencoded/uncompressed data. Lossy data compression is also known as irreversible or noisy compression. Entropy is defined as the quantity of information in a given set of data. Thus, one obvious advantage of lossy data compression is that the compression ratios can be larger than the entropy limit, all at the expense of information content. Many lossy data compression techniques seek to exploit various traits within the human senses to eliminate otherwise imperceptible data. For example, lossy data compression of visual imagery might seek to delete information content in excess of the display resolution or contrast ratio.

On the other hand, lossless data compression techniques provide an exact representation of the original uncompressed data. Simply stated, the decoded (or reconstructed) data is identical to the original unencoded/uncompressed data. Lossless data compression is also known as reversible or noiseless compression. Thus, lossless data compression has, as its current limit, a minimum representation defined by the entropy of a given data set.

There are various problems associated with the use of lossless compression techniques. One fundamental problem encountered with most lossless data compression techniques are their content sensitive behavior. This is often referred to as data dependency. Data dependency implies that the compression ratio achieved is highly contingent upon the content of the data being compressed. For example, database files often have large unused fields and high data redundancies, offering the opportunity to losslessly compress data at ratios of 5 to 1 or more. In contrast, concise software

programs have little to no data redundancy and, typically, will not losslessly compress better than 2 to 1.

Another problem with lossless compression is that there are significant variations in the compression ratio obtained when using a single lossless data compression technique for data streams having different data content and data size. This process is known as natural variation.

A further problem is that negative compression may occur when certain data compression techniques act upon many types of highly compressed data. Highly compressed data appears random and many data compression techniques will substantially expand, not compress this type of data.

For a given application, there are many factors that govern the applicability of various data compression techniques. These factors include compression ratio, encoding and decoding processing requirements, encoding and decoding time delays, compatibility with existing standards, and implementation complexity and cost, along with the adaptability and robustness to variations in input data. A direct relationship exists in the current art between compression ratio and the amount and complexity of processing required. One of the limiting factors in most existing prior art lossless data compression techniques is the rate at which the encoding and decoding processes are performed. Hardware and software implementation tradeoffs are often dictated by encoder and decoder complexity along with cost.

Another problem associated with lossless compression methods is determining the optimal compression technique for a given set of input data and intended application. To combat this problem, there are many conventional *content dependent* techniques that may be utilized. For instance, file type descriptors are typically appended to file names to describe the application programs that normally act upon the data contained within the file. In this manner data types, data structures, and formats within a given file may be ascertained. Fundamental limitations with this *content dependent* technique include:

- (1) the extremely large number of application programs, some of which do not possess published or documented file formats, data structures, or data type descriptors;
- (2) the ability for any data compression supplier or consortium to acquire, store, and access the vast amounts of data required to identify known file descriptors and associated data types, data structures, and formats; and

(3) the rate at which new application programs are developed and the need to update file format data descriptions accordingly.

An alternative technique that approaches the problem of selecting an appropriate lossless data compression technique is disclosed, for example, in U.S. Patent No.

5,467,087 to Chu entitled "High Speed Lossless Data Compression System" ("Chu").

FIG. 1 illustrates an embodiment of this data compression and decompression technique.

Data compression 1 comprises two phases, a data pre-compression phase 2 and a data compression phase 3. Data decompression 4 of a compressed input data stream is also comprised of two phases, a data type retrieval phase 5 and a data decompression phase 6.

During the data compression process 1, the data pre-compressor 2 accepts an uncompressed data stream, identifies the data type of the input stream, and generates a data type identification signal. The data compressor 3 selects a data compression method from a preselected set of methods to compress the input data stream, with the intention of producing the best available compression ratio for that particular data type.

There are several limitations associated with the Chu method. One such limitation is the need to unambiguously identify various data types. While these might include such common data types as ASCII, binary, or unicode, there, in fact, exists a broad universe of data types that fall outside the three most common data types. Examples of these alternate data types include: signed and unsigned integers of various lengths, differing types and precision of floating point numbers, pointers, other forms of character text, and a multitude of user defined data types. Additionally, data types may be interspersed or partially compressed, making data type recognition difficult and/or impractical. Another limitation is that given a known data type, or mix of data types within a specific set or subset of input data, it may be difficult and/or impractical to predict which data encoding technique yields the highest compression ratio.

Accordingly, there is a need for a data compression system and method that would address limitations in conventional data compression techniques as described above.

### **SUMMARY OF THE INVENTION**

The present invention is directed to systems and methods for providing fast and efficient data compression using a combination of content independent data compression and content dependent data compression. In one aspect of the invention, a method for

compressing data comprises the steps of:

analyzing a data block of an input data stream to identify a data type of the data block, the input data stream comprising a plurality of disparate data types;

performing content dependent data compression on the data block, if the data type  
5 of the data block is identified;

performing content independent data compression on the data block, if the data type of the data block is not identified.

In another aspect, the step of performing content independent data compression comprises: encoding the data block with a plurality of encoders to provide a plurality of  
10 encoded data blocks; determining a compression ratio obtained for each of the encoders; comparing each of the determined compression ratios with a first compression threshold; selecting for output the input data block and appending a null compression descriptor to the input data block, if all of the encoder compression ratios do not meet the first compression threshold; and selecting for output the encoded data block having the  
15 highest compression ratio and appending a corresponding compression type descriptor to the selected encoded data block, if at least one of the compression ratios meet the first compression threshold.

In another aspect, the step of performing content dependent compression comprises the steps of: selecting one or more encoders associated with the identified data  
20 type and encoding the data block with the selected encoders to provide a plurality of encoded data blocks; determining a compression ratio obtained for each of the selected encoders; comparing each of the determined compression ratios with a second compression threshold; selecting for output the input data block and appending a null compression descriptor to the input data block, if all of the encoder compression do not  
25 meet the second compression threshold; and selecting for output the encoded data block having the highest compression ratio and appending a corresponding compression type descriptor to the selected encoded data block, if at least one of the compression ratios meet the second compression threshold.

In yet another aspect, the step of performing content independent data  
30 compression on the data block, if the data type of the data block is not identified, comprises the steps of: estimating a desirability of using of one or more encoder types based one characteristics of the data block; and compressing the data block using one or

more desirable encoders.

In another aspect, the step of performing content dependent data compression on the data block, if the data type of the data block is identified, comprises the steps of:  
estimating a desirability of using of one or more encoder types based on characteristics of  
5 the data block; and compressing the data block using one or more desirable encoders.

In another aspect, the step of analyzing the data block comprises analyzing the data block to recognize one of a data type, data structure, data block format, file substructure, and/or file types. A further step comprises maintaining an association between encoder types and data types, data structures, data block formats, file  
10 substructure, and/or file types.

In yet another aspect of the invention, a method for compressing data comprises the steps of:

analyzing a data block of an input data stream to identify a data type of the data block, the input data stream comprising a plurality of disparate data types;  
15 performing content dependent data compression on the data block, if the data type of the data block is identified;  
determining a compression ratio of the compressed data block obtained using the content dependent compression and comparing the compression ratio with a first compression threshold; and  
20 performing content independent data compression on the data block, if the data type of the data block is not identified or if the compression ratio of the compressed data block obtained using the content dependent compression does not meet the first compression threshold.

Advantageously, the present invention employs a plurality of encoders applying a  
25 plurality of compression techniques on an input data stream so as to achieve maximum compression in accordance with the real-time or pseudo real-time data rate constraint. Thus, the output bit rate is not fixed and the amount, if any, of permissible data quality degradation is user or data specified.

These and other aspects, features and advantages of the present invention will  
30 become apparent from the following detailed description of preferred embodiments, which is to be read in connection with the accompanying drawings.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is a block/flow diagram of a content dependent high-speed lossless data compression and decompression system/method according to the prior art;

FIG. 2 is a block diagram of a content independent data compression system  
5 according to one embodiment of the present invention;

FIGs. 3a and 3b comprise a flow diagram of a data compression method according to one aspect of the present invention, which illustrates the operation of the data compression system of FIG. 2;

FIG. 4 is a block diagram of a content independent data compression system  
10 according to another embodiment of the present invention having an enhanced metric for selecting an optimal encoding technique;

FIGs. 5a and 5b comprise a flow diagram of a data compression method according to another aspect of the present invention, which illustrates the operation of the data compression system of FIG. 4;

FIG. 6 is a block diagram of a content independent data compression system  
15 according to another embodiment of the present invention having an *a priori* specified timer that provides real-time or pseudo real-time of output data;

FIGs. 7a and 7b comprise a flow diagram of a data compression method according to another aspect of the present invention, which illustrates the operation of the  
20 data compression system of FIG. 6;

FIG. 8 is a block diagram of a content independent data compression system according to another embodiment having an *a priori* specified timer that provides real-time or pseudo real-time of output data and an enhanced metric for selecting an optimal encoding technique;

FIG. 9 is a block diagram of a content independent data compression system  
25 according to another embodiment of the present invention having an encoding architecture comprising a plurality of sets of serially cascaded encoders;

FIGs. 10a and 10b comprise a flow diagram of a data compression method according to another aspect of the present invention, which illustrates the operation of the  
30 data compression system of FIG. 9;

FIG. 11 is block diagram of a content independent data decompression system according to one embodiment of the present invention;

FIG. 12 is a flow diagram of a data decompression method according to one aspect of the present invention, which illustrates the operation of the data compression system of FIG. 11;

FIGs. 13a and 13b comprise a block diagram of a data compression system comprising content dependent and content independent data compression, according to an embodiment of the present invention;

FIGs. 14a-14d comprise a flow diagram of a data compression method using both content dependent and content independent data compression, according to one aspect of the present invention;

FIGs. 15a and 15b comprise a block diagram of a data compression system comprising content dependent and content independent data compression, according to another embodiment of the present invention;

FIGs. 16a-16d comprise a flow diagram of a data compression method using both content dependent and content independent data compression, according to another aspect of the present invention;

FIGs. 17a and 17b comprise a block diagram of a data compression system comprising content dependent and content independent data compression, according to another embodiment of the present invention; and

FIGs. 18a-18d comprise a flow diagram of a data compression method using both content dependent and content independent data compression, according to another aspect of the present invention.

### **DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS**

The present invention is directed to systems and methods for providing data compression and decompression using content independent and content dependent data compression and decompression. In the following description, it is to be understood that system elements having equivalent or similar functionality are designated with the same reference numerals in the Figures. It is to be further understood that the present invention may be implemented in various forms of hardware, software, firmware, or a combination thereof. In particular, the system modules described herein are preferably implemented in software as an application program that is executable by, e.g., a general purpose computer or any machine or device having any suitable and preferred microprocessor



architecture. Preferably, the present invention is implemented on a computer platform including hardware such as one or more central processing units (CPU), a random access memory (RAM), and input/output (I/O) interface(s). The computer platform also includes an operating system and microinstruction code. The various processes and functions described herein may be either part of the microinstruction code or application programs which are executed via the operating system. In addition, various other peripheral devices may be connected to the computer platform such as an additional data storage device and a printing device.

It is to be further understood that, because some of the constituent system components described herein are preferably implemented as software modules, the actual system connections shown in the Figures may differ depending upon the manner in which the systems are programmed. It is to be appreciated that special purpose microprocessors may be employed to implement the present invention. Given the teachings herein, one of ordinary skill in the related art will be able to contemplate these and similar implementations or configurations of the present invention.

Referring now to FIG. 2 a block diagram illustrates a content independent data compression system according to one embodiment of the present invention. The data compression system includes a counter module 10 that receives as input an uncompressed or compressed data stream. It is to be understood that the system processes the input data stream in data blocks that may range in size from individual bits through complete files or collections of multiple files. Additionally, the data block size may be fixed or variable. The counter module 10 counts the size of each input data block (i.e., the data block size is counted in bits, bytes, words, any convenient data multiple or metric, or any combination thereof).

An input data buffer 20, operatively connected to the counter module 10, may be provided for buffering the input data stream in order to output an uncompressed data stream in the event that, as discussed in further detail below, every encoder fails to achieve a level of compression that exceeds an *a priori* specified minimum compression ratio threshold. It is to be understood that the input data buffer 20 is not required for implementing the present invention.

An encoder module 30 is operatively connected to the buffer 20 and comprises a set of encoders E1, E2, E3 ... En. The encoder set E1, E2, E3 ... En may include any

number "n" of those lossless encoding techniques currently well known within the art such as run length, Huffman, Lempel-Ziv Dictionary Compression, arithmetic coding, data compaction, and data null suppression. It is to be understood that the encoding techniques are selected based upon their ability to effectively encode different types of input data. It is to be appreciated that a full complement of encoders are preferably selected to provide a broad coverage of existing and future data types.

The encoder module 30 successively receives as input each of the buffered input data blocks (or unbuffered input data blocks from the counter module 10). Data compression is performed by the encoder module 30 wherein each of the encoders E1 .... En processes a given input data block and outputs a corresponding set of encoded data blocks. It is to be appreciated that the system affords a user the option to enable/disable any one or more of the encoders E1.... En prior to operation. As is understood by those skilled in the art, such feature allows the user to tailor the operation of the data compression system for specific applications. It is to be further appreciated that the encoding process may be performed either in parallel or sequentially. In particular, the encoders E1 through En of encoder module 30 may operate in parallel (i.e., simultaneously processing a given input data block by utilizing task multiplexing on a single central processor, via dedicated hardware, by executing on a plurality of processor or dedicated hardware systems, or any combination thereof). In addition, encoders E1 through En may operate sequentially on a given unbuffered or buffered input data block. This process is intended to eliminate the complexity and additional processing overhead associated with multiplexing concurrent encoding techniques on a single central processor and/or dedicated hardware, set of central processors and/or dedicated hardware, or any achievable combination. It is to be further appreciated that encoders of the identical type may be applied in parallel to enhance encoding speed. For instance, encoder E1 may comprise two parallel Huffman encoders for parallel processing of an input data block.

A buffer/counter module 40 is operatively connected to the encoding module 30 for buffering and counting the size of each of the encoded data blocks output from encoder module 30. Specifically, the buffer/counter 30 comprises a plurality of buffer/counters BC1, BC2, BC3 .... BCn, each operatively associated with a corresponding one of the encoders E1...En. A compression ratio module 50, operatively

connected to the output buffer/counter 40, determines the compression ratio obtained for each of the enabled encoders E1...En by taking the ratio of the size of the input data block to the size of the output data block stored in the corresponding buffer/counters BC1 ... BCn. In addition, the compression ratio module 50 compares each compression ratio with an *a priori*-specified compression ratio threshold limit to determine if at least one of the encoded data blocks output from the enabled encoders E1...En achieves a compression that exceeds an *a priori*-specified threshold. As is understood by those skilled in the art, the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. A description module 60, operatively coupled to the compression ratio module 50, appends a corresponding compression type descriptor to each encoded data block which is selected for output so as to indicate the type of compression format of the encoded data block.

The operation of the data compression system of FIG. 2 will now be discussed in further detail with reference to the flow diagram of FIGs. 3a and 3b. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the stream is received (step 300). As stated above, data compression is performed on a per data block basis. Accordingly, the first input data block in the input data stream is input into the counter module 10 that counts the size of the data block (step 302). The data block is then stored in the buffer 20 (step 304). The data block is then sent to the encoder module 30 and compressed by each (enabled) encoder E1 ... En (step 306). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder E1...En and maintained in a corresponding buffer (step 308), and the encoded data block size is counted (step 310).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10) to the size of each encoded data block output from the enabled encoders (step 312). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 314). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that notwithstanding that the current limit for lossless data compression is the entropy limit (the present

definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 316). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 316), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 318). A null data compression type descriptor is defined as any recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 320).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 316), then the encoded data block having the greatest compression ratio is selected (step 322). An appropriate data compression type descriptor is then appended (step 324).

A data compression type descriptor is defined as any recognizable data token or descriptor that indicates which data encoding technique has been applied to the data. It is to be understood that, since encoders of the identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 326).

After the encoded data block or the unencoded input data block is output (steps 326 and 320), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 328). If the input data stream includes additional data blocks (affirmative result in step 328), the next successive data block is received (step 330), its block size is counted (return to step 302) and the data compression process in repeated. This process is iterated for each data block in the input

data stream. Once the final input data block is processed (negative result in step 328), data compression of the input data stream is finished (step 322).

Since a multitude of data types may be present within a given input data block, it is often difficult and/or impractical to predict the level of compression that will be achieved by a specific encoder. Consequently, by processing the input data blocks with a plurality of encoding techniques and comparing the compression results, content free data compression is advantageously achieved. It is to be appreciated that this approach is scalable through future generations of processors, dedicated hardware, and software. As processing capacity increases and costs reduce, the benefits provided by the present invention will continue to increase. It should again be noted that the present invention may employ any lossless data encoding technique.

Referring now to Fig. 4, a block diagram illustrates a content independent data compression system according to another embodiment of the present invention. The data compression system depicted in FIG. 4 is similar to the data compression system of FIG. 2 except that the embodiment of FIG. 4 includes an enhanced metric functionality for selecting an optimal encoding technique. In particular, each of the encoders E1...En in the encoder module 30 is tagged with a corresponding one of user-selected encoder desirability factors 70. Encoder desirability is defined as an *a priori* user specified factor that takes into account any number of user considerations including, but not limited to, compatibility of the encoded data with existing standards, data error robustness, or any other aggregation of factors that the user wishes to consider for a particular application. Each encoded data block output from the encoder module 30 has a corresponding desirability factor appended thereto. A figure of merit module 80, operatively coupled to the compression ratio module 50 and the descriptor module 60, is provided for calculating a figure of merit for each of the encoded data blocks which possess a compression ratio greater than the compression ratio threshold limit. The figure of merit for each encoded data block is comprised of a weighted average of the *a priori* user specified threshold and the corresponding encoder desirability factor. As discussed below in further detail with reference to FIGs. 5a and 5b, the figure of merit substitutes the *a priori* user compression threshold limit for selecting and outputting encoded data blocks.

The operation of the data compression system of Fig. 4 will now be discussed in further detail with reference to the flow diagram of FIGs. 5a and 5b. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the stream is received (step 500). The size of the first data block is then determined by the counter module 10 (step 502). The data block is then stored in the buffer 20 (step 504). The data block is then sent to the encoder module 30 and compressed by each (enabled) encoder in the encoder set E1 ... En (step 506). Each encoded data block processed in the encoder module 30 is tagged with an encoder desirability factor that corresponds the particular encoding technique applied to the encoded data block (step 508). Upon completion of the encoding of the input data block, an encoded data block with its corresponding desirability factor is output from each (enabled) encoder E1...En and maintained in a corresponding buffer (step 510), and the encoded data block size is counted (step 512).

Next, a compression ratio obtained by each enabled encoder is calculated by taking the ratio of the size of the input data block (as determined by the input counter 10) to the size of the encoded data block output from each enabled encoder (step 514). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 516). A determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 518). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 518), then the original unencoded input data block is selected for output and a null data compression type descriptor (as discussed above) is appended thereto (step 520). Accordingly, the original unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 522).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 518), then a figure of merit is calculated for each encoded data block having a compression ratio which exceeds the compression ratio threshold limit (step 524). Again, the figure of merit for a given encoded data block is comprised of a weighted average of the *a priori* user specified threshold and the corresponding encoder desirability factor associated with the encoded data block. Next, the encoded data block having the greatest

figure of merit is selected for output (step 526). An appropriate data compression type descriptor is then appended (step 528) to indicate the data encoding technique applied to the encoded data block. The encoded data block (which has the greatest figure of merit) along with its corresponding data compression type descriptor is then output for  
5 subsequent data processing, storage, or transmittal (step 530).

After the encoded data block or the unencoded input data block is output (steps 530 and 522), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 532). If the input data stream includes additional data blocks (affirmative result in step 532), then the next successive data block  
10 is received (step 534), its block size is counted (return to step 502) and the data compression process is iterated for each successive data block in the input data stream. Once the final input data block is processed (negative result in step 532), data compression of the input data stream is finished (step 536).

Referring now to FIG. 6, a block diagram illustrates a data compression system  
15 according to another embodiment of the present invention. The data compression system depicted in FIG. 6 is similar to the data compression system discussed in detail above with reference to FIG. 2 except that the embodiment of FIG. 6 includes an *a priori* specified timer that provides real-time or pseudo real-time output data. In particular, an interval timer 90, operatively coupled to the encoder module 30, is preloaded with a user  
20 specified time value. The role of the interval timer (as will be explained in greater detail below with reference to FIGs. 7a and 7b) is to limit the processing time for each input data block processed by the encoder module 30 so as to ensure that the real-time, pseudo real-time, or other time critical nature of the data compression processes is preserved.

The operation of the data compression system of Fig. 6 will now be discussed in  
25 further detail with reference to the flow diagram of FIGs. 7a and 7b. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the data stream is received (step 700), and its size is determined by the counter module 10 (step 702). The data block is then stored in buffer 20 (step 704).

Next, concurrent with the completion of the receipt and counting of the first data  
30 block, the interval timer 90 is initialized (step 706) and starts counting towards a user-specified time limit. The input data block is then sent to the encoder module 30 wherein data compression of the data block by each (enabled) encoder E1 ... En commences (step

708). Next, a determination is made as to whether the user specified time expires before the completion of the encoding process (steps 710 and 712). If the encoding process is completed before or at the expiration of the timer, i.e., each encoder (E1 through En) completes its respective encoding process (negative result in step 710 and affirmative  
5 result in step 712), then an encoded data block is output from each (enabled) encoder E1...En and maintained in a corresponding buffer (step 714).

On the other hand, if the timer expires (affirmative result in 710), the encoding process is halted (step 716). Then, encoded data blocks from only those enabled encoders E1...En that have completed the encoding process are selected and maintained  
10 in buffers (step 718). It is to be appreciated that it is not necessary (or in some cases desirable) that some or all of the encoders complete the encoding process before the interval timer expires. Specifically, due to encoder data dependency and natural variation, it is possible that certain encoders may not operate quickly enough and, therefore, do not comply with the timing constraints of the end use. Accordingly, the  
15 time limit ensures that the real-time or pseudo real-time nature of the data encoding is preserved.

After the encoded data blocks are buffered (step 714 or 718), the size of each encoded data block is counted (step 720). Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as  
20 determined by the input counter 10) to the size of the encoded data block output from each enabled encoder (step 722). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 724). A determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 726). If there are no encoded data blocks having a compression ratio  
25 that exceeds the compression ratio threshold limit (negative determination in step 726), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 728). The original unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 730).

30 On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 726), then the encoded data block having the greatest compression ratio is selected



(step 732). An appropriate data compression type descriptor is then appended (step 734). The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 736).

5           After the encoded data block or the unencoded input data block is output (steps 730 or 736), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 738). If the input data stream includes additional data blocks (affirmative result in step 738), the next successive data block is received (step 740), its block size is counted (return to step 702) and the data  
10   compression process is repeated. This process is iterated for each data block in the input data stream, with each data block being processed within the user-specified time limit as discussed above. Once the final input data block is processed (negative result in step 738), data compression of the input data stream is complete (step 742).

Referring now to FIG. 8, a block diagram illustrates a content independent data  
15   compression system according to another embodiment of the present system. The data compression system of FIG. 8 incorporates all of the features discussed above in connection with the system embodiments of FIGs. 2, 4, and 6. For example, the system of FIG. 8 incorporates both the *a priori* specified timer for providing real-time or pseudo real-time of output data, as well as the enhanced metric for selecting an optimal encoding  
20   technique. Based on the foregoing discussion, the operation of the system of FIG. 8 is understood by those skilled in the art.

Referring now to FIG. 9, a block diagram illustrates a data compression system according to a preferred embodiment of the present invention. The system of FIG. 9 contains many of the features of the previous embodiments discussed above. However,  
25   this embodiment advantageously includes a cascaded encoder module 30c having an encoding architecture comprising a plurality of sets of serially-cascaded encoders  $E_{m,n}$ , where "m" refers to the encoding path (i.e., the encoder set) and where "n" refers to the number of encoders in the respective path. It is to be understood that each set of serially cascaded encoders can include any number of disparate and/or similar encoders (i.e., n  
30   can be any value for a given path m).

The system of FIG. 9 also includes a output buffer module 40c which comprises a plurality of buffer/counters B/C  $m,n$ , each associated with a corresponding one of the

encoders  $E_{m,n}$ . In this embodiment, an input data block is sequentially applied to successive encoders (encoder stages) in the encoder path so as to increase the data compression ratio. For example, the output data block from a first encoder  $E_{1,1}$ , is buffered and counted in  $B/C_{1,1}$ , for subsequent processing by a second encoder  $E_{1,2}$ .

- 5 Advantageously, these parallel sets of sequential encoders are applied to the input data stream to effect content free lossless data compression. This embodiment provides for multi-stage sequential encoding of data with the maximum number of encoding steps subject to the available real-time, pseudo real-time, or other timing constraints.

- As with each previously discussed embodiment, the encoders  $E_{m,n}$  may include
- 10 those lossless encoding techniques currently well known within the art, including: run length, Huffman, Lempel-Ziv Dictionary Compression, arithmetic coding, data compaction, and data null suppression. Encoding techniques are selected based upon their ability to effectively encode different types of input data. A full complement of encoders provides for broad coverage of existing and future data types. The input data
- 15 blocks may be applied simultaneously to the encoder paths (i.e., the encoder paths may operate in parallel, utilizing task multiplexing on a single central processor, or via dedicated hardware, or by executing on a plurality of processor or dedicated hardware systems, or any combination thereof). In addition, an input data block may be sequentially applied to the encoder paths. Moreover, each serially cascaded encoder path
- 20 may comprise a fixed (predetermined) sequence of encoders or a random sequence of encoders. Advantageously, by simultaneously or sequentially processing input data blocks via a plurality of sets of serially cascaded encoders, content free data compression is achieved.

- The operation of the data compression system of FIG. 9 will now be discussed in
- 25 further detail with reference to the flow diagram of FIGs. 10a and 10b. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the data stream is received (step 100), and its size is determined by the counter module 10 (step 102). The data block is then stored in buffer 20 (step 104).

- Next, concurrent with the completion of the receipt and counting of the first data
- 30 block, the interval timer 90 is initialized (step 106) and starts counting towards a user-specified time limit. The input data block is then sent to the cascade encoder module 30C wherein the input data block is applied to the first encoder (i.e., first encoding stage) in

each of the cascaded encoder paths  $E_{1,1} \dots E_{m,1}$  (step 108). Next, a determination is made as to whether the user specified time expires before the completion of the first stage encoding process (steps 110 and 112). If the first stage encoding process is completed before the expiration of the timer, i.e., each encoder ( $E_{1,1} \dots E_{m,1}$ ) completes its  
5 respective encoding process (negative result in step 110 and affirmative result in step 112), then an encoded data block is output from each encoder  $E_{1,1} \dots E_{m,1}$  and maintained in a corresponding buffer (step 114). Then for each cascade encoder path, the output of the completed encoding stage is applied to the next successive encoding stage in the cascade path (step 116). This process (steps 110, 112, 114, and 116) is repeated  
10 until the earlier of the timer expiration (affirmative result in step 110) or the completion of encoding by each encoder stage in the serially cascaded paths, at which time the encoding process is halted (step 118).

Then, for each cascade encoder path, the buffered encoded data block output by the last encoder stage that completes the encoding process before the expiration of the  
15 timer is selected for further processing (step 120). Advantageously, the interim stages of the multi-stage data encoding process are preserved. For example, the results of encoder  $E_{1,1}$  are preserved even after encoder  $E_{1,2}$  begins encoding the output of encoder  $E_{1,1}$ . If the interval timer expires after encoder  $E_{1,1}$  completes its respective encoding process but before encoder  $E_{1,2}$  completes its respective encoding process, the encoded data  
20 block from encoder  $E_{1,1}$  is complete and is utilized for calculating the compression ratio for the corresponding encoder path. The incomplete encoded data block from encoder  $E_{1,2}$  is either discarded or ignored.

It is to be appreciated that it is not necessary (or in some cases desirable) that some or all of the encoders in the cascade encoder paths complete the encoding process  
25 before the interval timer expires. Specifically, due to encoder data dependency, natural variation and the sequential application of the cascaded encoders, it is possible that certain encoders may not operate quickly enough and therefore do not comply with the timing constraints of the end use. Accordingly, the time limit ensures that the real-time or pseudo real-time nature of the data encoding is preserved.

30 After the encoded data blocks are selected (step 120), the size of each encoded data block is counted (step 122). Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined

by the input counter 10) to the size of the encoded data block output from each encoder (step 124). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 126). A determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 128). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 128), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 130). The original unencoded data block and its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 132).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 128), then a figure of merit is calculated for each encoded data block having a compression ratio which exceeds the compression ratio threshold limit (step 134). Again, the figure of merit for a given encoded data block is comprised of a weighted average of the *a priori* user specified threshold and the corresponding encoder desirability factor associated with the encoded data block. Next, the encoded data block having the greatest figure of merit is selected (step 136). An appropriate data compression type descriptor is then appended (step 138) to indicate the data encoding technique applied to the encoded data block. For instance, the data type compression descriptor can indicate that the encoded data block was processed by either a single encoding type, a plurality of sequential encoding types, and a plurality of random encoding types. The encoded data block (which has the greatest figure of merit) along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 140).

After the unencoded data block or the encoded data input data block is output (steps 132 and 140), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 142). If the input data stream includes additional data blocks (affirmative result in step 142), then the next successive data block is received (step 144), its block size is counted (return to step 102) and the data compression process is iterated for each successive data block in the input data stream.

Once the final input data block is processed (negative result in step 142), data compression of the input data stream is finished (step 146).

Referring now to FIG. 11, a block diagram illustrates a data decompression system according to one embodiment of the present invention. The data decompression system preferably includes an input buffer 1100 that receives as input an uncompressed or compressed data stream comprising one or more data blocks. The data blocks may range in size from individual bits through complete files or collections of multiple files. Additionally, the data block size may be fixed or variable. The input data buffer 1100 is preferably included (not required) to provide storage of input data for various hardware implementations. A descriptor extraction module 1102 receives the buffered (or unbuffered) input data block and then parses, lexically, syntactically, or otherwise analyzes the input data block using methods known by those skilled in the art to extract the data compression type descriptor associated with the data block. The data compression type descriptor may possess values corresponding to null (no encoding applied), a single applied encoding technique, or multiple encoding techniques applied in a specific or random order (in accordance with the data compression system embodiments and methods discussed above).

A decoder module 1104 includes a plurality of decoders D1...Dn for decoding the input data block using a decoder, set of decoders, or a sequential set of decoders corresponding to the extracted compression type descriptor. The decoders D1...Dn may include those lossless encoding techniques currently well known within the art, including: run length, Huffman, Lempel-Ziv Dictionary Compression, arithmetic coding, data compaction, and data null suppression. Decoding techniques are selected based upon their ability to effectively decode the various different types of encoded input data generated by the data compression systems described above or originating from any other desired source. As with the data compression systems discussed above, the decoder module 1104 may include multiple decoders of the same type applied in parallel so as to reduce the data decoding time.

The data decompression system also includes an output data buffer 1106 for buffering the decoded data block output from the decoder module 1104.

The operation of the data decompression system of FIG. 11 will be discussed in further detail with reference to the flow diagram of FIG. 12. A data stream comprising one or more data blocks of compressed or uncompressed data is input into the data decompression system and the first data block in the stream is received (step 1200) and maintained in the buffer (step 1202). As  
5 with the data compression systems discussed above, data decompression is performed on a per data block basis. The data compression type descriptor is then extracted from the input data block (step 1204). A determination is then made as to whether the data compression type descriptor is null (step 1206). If the data compression type descriptor is determined to be null (affirmative result in step 1206), then no decoding is applied to the input data block and the  
10 original undecoded data block is output (or maintained in the output buffer) (step 1208).

On the other hand, if the data compression type descriptor is determined to be any value other than null (negative result in step 1206), the corresponding decoder or decoders are then selected (step 1210) from the available set of decoders D1...Dn in the decoding module 1104. It is to be understood that the data compression type descriptor may mandate the application of: a  
15 single specific decoder, an ordered sequence of specific decoders, a random order of specific decoders, a class or family of decoders, a mandatory or optional application of parallel decoders, or any combination or permutation thereof. The input data block is then decoded using the selected decoders (step 1212), and output (or maintained in the output buffer 1106) for subsequent data processing, storage, or transmittal (step 1214). A determination is then made as  
20 to whether the input data stream contains additional data blocks to be processed (step 1216). If the input data stream includes additional data blocks (affirmative result in step 1216), the next successive data block is received (step 1220), and buffered (return to step 1202). Thereafter, the data decompression process is iterated for each data block in the input data stream. Once the final input data block is processed (negative result in step 1216), data decompression of the input  
25 data stream is finished (step 1218).

In other embodiments of the present invention described below, data compression is achieved using a combination of content dependent data compression and content independent data compression. For example, FIGs. 13a and 13b are block diagrams illustrating a data compression system employing both content independent and content dependent data  
30 compression according to one embodiment of the present invention, wherein content independent

data compression is applied to a data block when the content of the data block cannot be identified or is not associable with a specific data compression algorithm. The data compression system comprises a counter module 10 that receives as input an uncompressed or compressed data stream. It is to be understood that the system processes the input data stream in data blocks that may range in size from individual bits through complete files or collections of multiple files. Additionally, the data block size may be fixed or variable. The counter module 10 counts the size of each input data block (i.e., the data block size is counted in bits, bytes, words, any convenient data multiple or metric, or any combination thereof).

An input data buffer 20, operatively connected to the counter module 10, may be provided for buffering the input data stream in order to output an uncompressed data stream in the event that, as discussed in further detail below, every encoder fails to achieve a level of compression that exceeds *a priori* specified content independent or content dependent minimum compression ratio thresholds. It is to be understood that the input data buffer 20 is not required for implementing the present invention.

A content dependent data recognition module 1300 analyzes the incoming data stream to recognize data types, data structures, data block formats, file substructures, file types, and/or any other parameters that may be indicative of either the data type/content of a given data block or the appropriate data compression algorithm or algorithms (in serial or in parallel) to be applied. Optionally, a data file recognition list(s) or algorithm(s) 1310 module may be employed to hold and/or determine associations between recognized data parameters and appropriate algorithms. Each data block that is recognized by the content data compression module 1300 is routed to a content dependent encoder module 1320, if not the data is routed to the content independent encoder module 30.

A content dependent encoder module 1320 is operatively connected to the content dependent data recognition module 1300 and comprises a set of encoders D1, D2, D3 ... Dm. The encoder set D1, D2, D3 ... Dm may include any number "n" of those lossless or lossy encoding techniques currently well known within the art such as MPEG4, various voice codecs, MPEG3, AC3, AAC, as well as lossless algorithms such as run length, Huffman, Lempel-Ziv Dictionary Compression, arithmetic coding, data compaction, and data null suppression. It is to be understood that the encoding techniques are selected based upon their ability to effectively

encode different types of input data. It is to be appreciated that a full complement of encoders and/or codecs are preferably selected to provide a broad coverage of existing and future data types.

5 The content independent encoder module 30, which is operatively connected to the content dependent data recognition module 1300, comprises a set of encoders E1, E2, E3 ... En. The encoder set E1, E2, E3 ... En may include any number "n" of those lossless encoding techniques currently well known within the art such as run length, Huffman, Lempel-Ziv Dictionary Compression, arithmetic coding, data compaction, and data null suppression. Again, it is to be understood that the encoding techniques are selected based upon their ability to  
10 effectively encode different types of input data. It is to be appreciated that a full complement of encoders are preferably selected to provide a broad coverage of existing and future data types.

The encoder modules (content dependent 1320 and content independent 30) selectively receive the buffered input data blocks (or unbuffered input data blocks from the counter module 10) from module 1300 based on the results of recognition. Data compression is performed by  
15 the respective encoder modules wherein some or all of the encoders D1....Dm or E1 .... En processes a given input data block and outputs a corresponding set of encoded data blocks. It is to be appreciated that the system affords a user the option to enable/disable any one or more of the encoders D1....Dm and E1.... En prior to operation. As is understood by those skilled in the art, such feature allows the user to tailor the operation of the data compression system for  
20 specific applications. It is to be further appreciated that the encoding process may be performed either in parallel or sequentially. In particular, the encoder set D1 through Dm of encoder module 1320 and/or the encoder set E1 through En of encoder module 30 may operate in parallel (i.e., simultaneously processing a given input data block by utilizing task multiplexing on a single central processor, via dedicated hardware, by executing on a plurality of processor or  
25 dedicated hardware systems, or any combination thereof). In addition, encoders D1 through Dm and E1 through En may operate sequentially on a given unbuffered or buffered input data block. This process is intended to eliminate the complexity and additional processing overhead associated with multiplexing concurrent encoding techniques on a single central processor and/or dedicated hardware, set of central processors and/or dedicated hardware, or any achievable  
30 combination. It is to be further appreciated that encoders of the identical type may be applied in



parallel to enhance encoding speed. For instance, encoder E1 may comprise two parallel Huffman encoders for parallel processing of an input data block. It should be further noted that one or more algorithms may be implemented in dedicated hardware such as an MPEG4 or MP3 encoding integrated circuit.

5           Buffer/counter modules 1330 and 40 are operatively connected to their respective encoding modules 1320 and 30, for buffering and counting the size of each of the encoded data blocks output from the respective encoder modules. Specifically, the content dependent buffer/counter 1330 comprises a plurality of buffer/counters BCD1, BCD2, BCD3 ....BCDm, each operatively associated with a corresponding one of the encoders D1...Dm. Similarly the  
10 content independent buffer/counters BCE1, BCE2, BCE3 ....BCEn, each operatively associated with a corresponding one of the encoders E1...En. A compression ratio module 1340, operatively connected to the content dependent output buffer/counters 1330 and content independent buffer/counters 40 determines the compression ratio obtained for each of the enabled encoders D1 ...Dm and or E1...En by taking the ratio of the size of the input data block  
15 to the size of the output data block stored in the corresponding buffer/counters BCD1, BCD2, BCD3 ....BCDm and or BCE1, BCE2, BCE3 ....BCEn. In addition, the compression ratio module 1340 compares each compression ratio with an *a priori*-specified compression ratio threshold limit to determine if at least one of the encoded data blocks output from the enabled encoders BCD1, BCD2, BCD3 ....BCDm and or BCE1, BCE2, BCE3 ....BCEn achieves a  
20 compression that meets an *a priori*-specified threshold. As is understood by those skilled in the art, the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It should be noted that different threshold values may be applied to content dependent and content independent encoded data. Further these thresholds may be adaptively modified based upon enabled encoders in either  
25 or both the content dependent or content independent encoder sets, along with any associated parameters. A compression type description module 1350, operatively coupled to the compression ratio module 1340, appends a corresponding compression type descriptor to each encoded data block which is selected for output so as to indicate the type of compression format of the encoded data block.

30           A mode of operation of the data compression system of FIGS. 13a and 13b will now be

discussed with reference to the flow diagrams of FIGs. 14a-14d, which illustrates a method for performing data compression using a combination of content dependent and content independent data compression. In general, content independent data compression is applied to a given data block when the content of a data block cannot be identified or is not associated with a specific data compression algorithm. More specifically, referring to FIG. 14a, a data stream comprising one or more data blocks is input into the data compression system and the first data block in the stream is received (step 1400). As stated above, data compression is performed on a per data block basis. As previously stated a data block may represent any quantity of data from a single bit through a multiplicity of files or packets and may vary from block to block. Accordingly, the first input data block in the input data stream is input into the counter module 10 that counts the size of the data block (step 1402). The data block is then stored in the buffer 20 (step 1404). The data block is then analyzed on a per block or multi-block basis by the content dependent data recognition module 1300 (step 1406). If the data stream content is not recognized utilizing the recognition list(s) or algorithms(s) module 1310 (step 1408) the data is routed to the content independent encoder module 30 and compressed by each (enabled) encoder E1 ... En (step 1410). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder E1 ... En and maintained in a corresponding buffer (step 1412), and the encoded data block size is counted (step 1414).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1416). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1418). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art. Additionally the content independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1420). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1420), then the original  
5 unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 1434). A null data compression type descriptor is defined as any recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal  
10 (step 1436).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 1420), then the encoded data block having the greatest compression ratio is selected (step 1422). An appropriate data compression type descriptor is then appended (step 1424). A data compression type  
15 descriptor is defined as any recognizable data token or descriptor that indicates which data encoding technique has been applied to the data. It is to be understood that, since encoders of the identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the  
20 greatest compression ratio along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1426).

As previously stated the data block stored in the buffer 20 (step 1404) is analyzed on a per block or multi-block basis by the content dependent data recognition module 1300 (step 1406). If the data stream content is recognized utilizing the recognition list(s) or algorithm(s)  
25 module 1310 (step 1434) the appropriate content dependent algorithms are enabled and initialized (step 1436), and the data is routed to the content dependent encoder module 1320 and compressed by each (enabled) encoder D1 ... Dm (step 1438). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder D1...Dm and maintained in a corresponding buffer (step 1440), and the encoded data block size is counted  
30 (step 1442).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1444). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1448). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that many of these algorithms may be lossy, and as such the limits may be subject to or modified by an end target storage, listening, or viewing device. Further notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art. Additionally the content independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1420). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1420), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 1434). A null data compression type descriptor is defined as any recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1436).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 1420), then the encoded data block having the greatest compression ratio is selected (step 1422). An appropriate data compression type descriptor is then appended (step 1424). A data compression type descriptor is defined as any recognizable data token or descriptor that indicates which data encoding technique has been applied to the data. It is to be understood that, since encoders of the

identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then  
5 output for subsequent data processing, storage, or transmittal (step 1426).

After the encoded data block or the unencoded data input data block is output (steps 1426 and 1436), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 1428). If the input data stream includes additional data blocks (affirmative result in step 1428), the next successive data block is received (step 1432), its block  
10 size is counted (return to step 1402) and the data compression process is repeated. This process is iterated for each data block in the input data stream. Once the final input data block is processed (negative result in step 1428), data compression of the input data stream is finished (step 1430).

Since a multitude of data types may be present within a given input data block, it is often  
15 difficult and/or impractical to predict the level of compression that will be achieved by a specific encoder. Consequently, by processing the input data blocks with a plurality of encoding techniques and comparing the compression results, content free data compression is advantageously achieved. Further the encoding may be lossy or lossless dependent upon the input data types. Further if the data type is not recognized the default content independent  
20 lossless compression is applied. It is not a requirement that this process be deterministic - in fact a certain probability may be applied if occasional data loss is permitted. It is to be appreciated that this approach is scalable through future generations of processors, dedicated hardware, and software. As processing capacity increases and costs reduce, the benefits provided by the present invention will continue to increase. It should again be noted that the present invention may  
25 employ any lossless data encoding technique.

FIGs. 15a and 15b are block diagrams illustrating a data compression system employing both content independent and content dependent data compression according to another embodiment of the present invention. The system in FIGs. 15a and 15b is similar in operation to the system of FIGs. 13a and 13b in that content independent data compression is applied to a  
30 data block when the content of the data block cannot be identified or is not associable with a

specific data compression algorithm. The system of FIGs. 15a and 15b additionally performs content independent data compression on a data block when the compression ratio obtained for the data block using the content dependent data compression does not meet a specified threshold.

A mode of operation of the data compression system of FIGs. 15a and 15b will now be  
5 discussed with reference to the flow diagram of FIGs. 16a-16d, which illustrates a method for performing data compression using a combination of content dependent and content independent data compression. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the stream is received (step 1600). As stated above, data compression is performed on a per data block basis. As previously stated a data  
10 block may represent any quantity of data from a single bit through a multiplicity of files or packets and may vary from block to block. Accordingly, the first input data block in the input data stream is input into the counter module 10 that counts the size of the data block (step 1602). The data block is then stored in the buffer 20 (step 1604). The data block is then analyzed on a per block or multi-block basis by the content dependent data recognition module 1300 (step  
15 1606). If the data stream content is not recognized utilizing the recognition list(s) or algorithms(s) module 1310 (Step 1608) the data is routed to the content independent encoder module 30 and compressed by each (enabled) encoder E1 ... En (step 1610). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder E1...En and maintained in a corresponding buffer (step 1612), and the encoded data  
20 block size is counted (step 1614).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1616). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1618). It is to be  
25 understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless  
30 data compression ratios beyond what is currently known within the art. Additionally the content

independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1620). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1620), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 1634). A null data compression type descriptor is defined as any recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1636).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 1620), then the encoded data block having the greatest compression ratio is selected (step 1622). An appropriate data compression type descriptor is then appended (step 1624). A data compression type descriptor is defined as any recognizable data token or descriptor that indicates which data encoding technique has been applied to the data. It is to be understood that, since encoders of the identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1626).

As previously stated the data block stored in the buffer 20 (step 1604) is analyzed on a per block or multi-block basis by the content dependent data recognition module 1300 (step 1606). If the data stream content is recognized utilizing the recognition list(s) or algorithm(s) module 1310 (step 1634) the appropriate content dependent algorithms are enabled and initialized (step 1636) and the data is routed to the content dependent encoder module 1620 and compressed by each (enabled) encoder D1 ... Dm (step 1638). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder D1 ... Dm

and maintained in a corresponding buffer (step 1640), and the encoded data block size is counted (step 1642).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1644). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1648). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that many of these algorithms may be lossy, and as such the limits may be subject to or modified by an end target storage, listening, or viewing device. Further notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art. Additionally the content independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1648). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1620), then the original unencoded input data block is routed to the content independent encoder module 30 and the process resumes with compression utilizing content independent encoders (step 1610).

After the encoded data block or the unencoded data input data block is output (steps 1626 and 1636), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 1628). If the input data stream includes additional data blocks (affirmative result in step 1628), the next successive data block is received (step 1632), its block size is counted (return to step 1602) and the data compression process is repeated. This process is iterated for each data block in the input data stream. Once the final input data block is processed (negative result in step 1628), data compression of the input data stream is finished (step 1630).



FIGs. 17a and 17b are block diagrams illustrating a data compression system employing both content independent and content dependent data compression according to another embodiment of the present invention. The system in FIGs. 17a and 17b is similar in operation to the system of FIGs. 13a and 13b in that content independent data compression is applied to a data block when the content of the data block cannot be identified or is not associable with a specific data compression algorithm. The system of FIGs. 17a and 17b additionally uses *a priori* estimation algorithms or look-up tables to estimate the desirability of using content independent data compression encoders and/or content dependent data compression encoders and selecting appropriate algorithms or subsets thereof based on such estimation.

More specifically, a content dependent data recognition and or estimation module 1700 is utilized to analyze the incoming data stream for recognition of data types, data structures, data block formats, file substructures, file types, or any other parameters that may be indicative of the appropriate data compression algorithm or algorithms ( in serial or in parallel) to be applied. Optionally, a data file recognition list(s) or algorithm(s) 1710 module may be employed to hold associations between recognized data parameters and appropriate algorithms. If the content data compression module recognizes a portion of the data, that portion is routed to the content dependent encoder module 1320, if not the data is routed to the content independent encoder module 30. It is to be appreciated that process of recognition (modules 1700 and 1710) is not limited to a deterministic recognition, but may further comprise a probabilistic estimation of which encoders to select for compression from the set of encoders of the content dependent module 1320 or the content independent module 30. For example, a method may be employed to compute statistics of a data block whereby a determination that the locality of repetition of characters in a data stream is determined is high can suggest a text document, which may be beneficially compressed with a lossless dictionary type algorithm. Further the statistics of repeated characters and relative frequencies may suggest a specific type of dictionary algorithm. Long strings will require a wide dictionary file while a wide diversity of strings may suggest a deep dictionary. Statistics may also be utilized in algorithms such as Huffman where various character statistics will dictate the choice of different Huffman compression tables. This technique is not limited to lossless algorithms but may be widely employed with lossy algorithms. Header information in frames for video files can imply a specific data resolution.

The estimator then may select the appropriate lossy compression algorithm and compression parameters (amount of resolution desired). As shown in previous embodiments of the present invention, desirability of various algorithms and now associated resolutions with lossy type algorithms may also be applied in the estimation selection process.

5       A mode of operation of the data compression system of FIGs. 17a and 17b will now be discussed with reference to the flow diagrams of FIGs. 18a-18d. The method of FIGs. 18a-18d use *a priori* estimation algorithms or look-up tables to estimate the desirability or probability of using content independent data compression encoders or content dependent data compression encoders, and select appropriate or desirable algorithms or subsets thereof based on such  
10       estimates. A data stream comprising one or more data blocks is input into the data compression system and the first data block in the stream is received (step 1800). As stated above, data compression is performed on a per data block basis. As previously stated a data block may represent any quantity of data from a single bit through a multiplicity of files or packets and may vary from block to block. Accordingly, the first input data block in the input data stream is input  
15       into the counter module 10 that counts the size of the data block (step 1802). The data block is then stored in the buffer 20 (step 1804). The data block is then analyzed on a per block or multi-block basis by the content dependent / content independent data recognition module 1700 (step 1806). If the data stream content is not recognized utilizing the recognition list(s) or algorithms(s) module 1710 (step 1808) the data is to the content independent encoder module  
20       30. An estimate of the best content independent encoders is performed (step 1850) and the appropriate encoders are enabled and initialized as applicable. The data is then compressed by each (enabled) encoder E1 ... En (step 1810). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder E1...En and maintained in a corresponding buffer (step 1812), and the encoded data block size is counted (step 1814).

25       Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1816). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1818). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no  
30       data compression or expansion, or any arbitrarily desired compression limit. It is to be further

understood that notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art. Additionally the content independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1820). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1820), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 1834). A null data compression type descriptor is defined as any recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1836).

On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 1820), then the encoded data block having the greatest compression ratio is selected (step 1822). An appropriate data compression type descriptor is then appended (step 1824). A data compression type descriptor is defined as any recognizable data token or descriptor that indicates which data encoding technique has been applied to the data. It is to be understood that, since encoders of the identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1826).

As previously stated the data block stored in the buffer 20 (step 1804) is analyzed on a per block or multi-block basis by the content dependent data recognition module 1300 (step 1806). If the data stream content is recognized or estimated utilizing the recognition list(s) or

algorithms(s) module 1710 (affirmative result in step 1808) the recognized data type/file or block is selected based on a list or algorithm (step 1838) and an estimate of the desirability of using the associated content dependent algorithms can be determined (step 1840). For instance, even though a recognized data type may be associated with three different encoders, an estimation of the desirability of using each encoder may result in only one or two of the encoders being actually selected for use. The data is routed to the content dependent encoder module 1320 and compressed by each (enabled) encoder D1 ... Dm (step 1842). Upon completion of the encoding of the input data block, an encoded data block is output from each (enabled) encoder D1...Dm and maintained in a corresponding buffer (step 1844), and the encoded data block size is counted (step 1846).

Next, a compression ratio is calculated for each encoded data block by taking the ratio of the size of the input data block (as determined by the input counter 10 to the size of each encoded data block output from the enabled encoders (step 1848). Each compression ratio is then compared with an *a priori*-specified compression ratio threshold (step 1850). It is to be understood that the threshold limit may be specified as any value inclusive of data expansion, no data compression or expansion, or any arbitrarily desired compression limit. It is to be further understood that many of these algorithms may be lossy, and as such the limits may be subject to or modified by an end target storage, listening, or viewing device. Further notwithstanding that the current limit for lossless data compression is the entropy limit (the present definition of information content) for the data, the present invention does not preclude the use of future developments in lossless data compression that may increase lossless data compression ratios beyond what is currently known within the art. Additionally the content independent data compression threshold may be different from the content dependent threshold and either may be modified by the specific enabled encoders.

After the compression ratios are compared with the threshold, a determination is made as to whether the compression ratio of at least one of the encoded data blocks exceeds the threshold limit (step 1820). If there are no encoded data blocks having a compression ratio that exceeds the compression ratio threshold limit (negative determination in step 1820), then the original unencoded input data block is selected for output and a null data compression type descriptor is appended thereto (step 1834). A null data compression type descriptor is defined as any

recognizable data token or descriptor that indicates no data encoding has been applied to the input data block. Accordingly, the unencoded input data block with its corresponding null data compression type descriptor is then output for subsequent data processing, storage, or transmittal (step 1836).

5        On the other hand, if one or more of the encoded data blocks possess a compression ratio greater than the compression ratio threshold limit (affirmative result in step 1820), then the encoded data block having the greatest compression ratio is selected (step 1822). An appropriate data compression type descriptor is then appended (step 1824). A data compression type descriptor is defined as any recognizable data token or descriptor that indicates which data  
10        encoding technique has been applied to the data. It is to be understood that, since encoders of the identical type may be applied in parallel to enhance encoding speed (as discussed above), the data compression type descriptor identifies the corresponding encoding technique applied to the encoded data block, not necessarily the specific encoder. The encoded data block having the greatest compression ratio along with its corresponding data compression type descriptor is then  
15        output for subsequent data processing, storage, or transmittal (step 1826).

After the encoded data block or the unencoded data input data block is output (steps 1826 and 1836), a determination is made as to whether the input data stream contains additional data blocks to be processed (step 1828). If the input data stream includes additional data blocks (affirmative result in step 1428), the next successive data block is received (step 1832), its block  
20        size is counted (return to step 1802) and the data compression process is repeated. This process is iterated for each data block in the input data stream. Once the final input data block is processed (negative result in step 1828), data compression of the input data stream is finished (step 1830).

It is to be appreciated that in the embodiments described above with reference to FIGS.  
25        13-18, an *a priori* specified time limit or any other real-time requirement may be employed to achieve practical and efficient real-time operation.

Although illustrative embodiments have been described herein with reference to the accompanying drawings, it is to be understood that the present invention is not limited to those precise embodiments, and that various other changes and modifications may be affected therein  
30        by one skilled in the art without departing from the scope or spirit of the invention. All such

changes and modifications are intended to be included within the scope of the invention as defined by the appended claims.